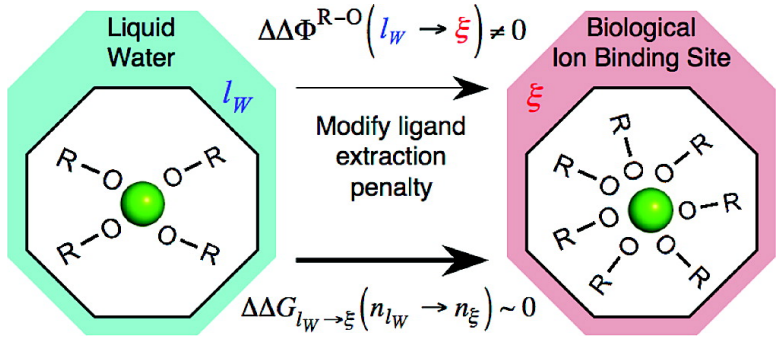
Article

Structural Transitions in Ion Coordination Driven by Changes in Competition for Ligand Binding

Sameer Varma, and Susan B. Rempe

J. Am. Chem. Soc., **2008**, 130 (46), 15405-15419 • DOI: 10.1021/ja803575y • Publication Date (Web): 28 October 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 10/28/2008

Structural Transitions in Ion Coordination Driven by Changes in Competition for Ligand Binding

Sameer Varma* and Susan B. Rempe*

Sandia National Laboratories, Albuquerque, New Mexico 87185

Received May 20, 2008; E-mail: svarma@sandia.gov; slrempe@sandia.gov

Abstract: Transferring Na^+ and K^+ ions from their preferred coordination states in water to states having different coordination numbers incurs a free energy cost. In several examples in nature, however, these ions readily partition from aqueous-phase coordination states into spatial regions having much higher coordination numbers. Here we utilize statistical theory of solutions, quantum chemical simulations, classical mechanics simulations, and structural informatics to understand this aspect of ion partitioning. Our studies lead to the identification of a specific role of the solvation environment in driving transitions in ion coordination structures. Although ion solvation in liquid media is an exergonic reaction overall, we find it is also associated with considerable free energy penalties for extracting ligands from their solvation environments to form coordinated ion complexes. Reducing these penalties increases the stabilities of higher-order coordinations and brings down the energetic cost to partition ions from water into overcoordinated binding sites in biomolecules. These penalties can be lowered via a reduction in direct favorable interactions of the coordinating ligands with all atoms other than the ions themselves. A significant reduction in these penalties can, in fact, also drive up ion coordination preferences. Similarly, an increase in these penalties can lower ion coordination preferences, akin to a Hofmeister effect. Since such structural transitions are effected by the properties of the solvation phase, we anticipate that they will also occur for other ions. The influence of other factors, including ligand density, ligand chemistry, and temperature, on the stabilities of ion coordination structures are also explored.

1. Introduction

Na^+ and K^+ ions play vital roles in numerous biochemical reactions as their binding affects activities of several globular proteins[1,2] and RNA[3] enzymes. In addition, their direct interactions with functional groups in ion channels enables regulation of their concentration gradients across cellular membranes, which allows a wide variety of high-level physiological tasks to be accomplished: from nutrient uptake and volume control in individual cells, to the generation and propagation of nerve signals and regulation of muscle activity such as in a heartbeat.[4] Direct interactions or chelation of these ions with biological molecules occurs only when they partially or completely shed the water molecules present within their inner coordination shells in the aqueous phase. Here we carry out a combination of quantum chemical simulations, classical molecular dynamics (MD) simulations, and structural informatics studies to understand the details of one specific aspect of this ion-partitioning phenomenon.

Recent studies based on neutron diffraction experiments[5−7] and *ab initio* molecular dynamics (AIMD) simulations[8−11] strongly suggest that in aqueous phase Na^+ and K^+ ions coordinate with five and six water molecules, respectively. These values correspond to the average numbers of water oxygens within the spherical subvolume $\Gamma = 4/3\pi r^3$ bounded by the principal minima of their radial distribution functions $g(r)$, that is, for $r \leq 3.0$ Å for Na^+ and $r \leq 3.5$ Å for K^+. Data from these AIMD simulations,[8−11] and also from longer time-scale trajectories generated as part of this work, show how coordination numbers, n, are distributed within Γ (Figure 1). Several noteworthy observations regarding these distributions can be made. First, the preferred coordination numbers $P_\Gamma(n)$ for subvolumes $\Gamma(r = 3.0)$ for Na^+ and $\Gamma(r = 3.5)$ for K^+ need not match the respective average coordination numbers of these ions. This results from asymmetric distribution of $P_\Gamma(n)$ around the most preferred values. Furthermore, the values of both preferred and average coordinations depend upon the size of Γ chosen for analysis. Finally and most importantly, the probability $P_\Gamma(n)$ of finding a coordination number n within Γ decreases sharply as one departs further away from the most preferred numbers. If the subvolume around the ion were chosen to be smaller, for example, where $r = 3.0$ Å for K^+, the preferred coordination number drops to four,[10−12] and the probability of

(1) Glusker, J. P. *Adv. Protein Chem.* **1991**, *42*, 1–76.
(2) Page, M. J.; Di Cera, E. *Physiol. Rev.* **2006**, *86*, 1049–1092.
(3) Draper, D. E.; Misra, V. K. *Nat. Struct. Biol.* **1998**, *5*, 927–930.
(4) Hille, B. *Ionic Channels of Excitable Membranes*, 3rd ed.; Sinauer Associates: Sunderland, MA, 2001; p 607.
(5) Neilson, G. W.; Mason, P. E.; Ramos, S.; Sullivan, D. *Philos. Trans. R. Soc. London, Ser. A* **2001**, *359*, 1575–1591.
(6) Ansell, S.; Barnes, A. C.; Mason, P. E.; Neilson, G. W.; Ramos, S. *Biophys. Chem.* **2006**, *124*, 171–179.
(7) Soper, A. K.; Weckstrom, K. *Biophys. Chem.* **2006**, *124*, 180–191.
(8) White, J. A.; Schwegler, E.; Galli, G.; Gygi, F. *J. Chem. Phys.* **2000**, *113*, 4668–4673.
(9) Rempe, S. B.; Pratt, L. R. *Fluid Phase Equilib.* **2001**, *183*, 121–132.
(10) Rempe, S. B.; Asthagiri, D.; Pratt, L. R. *Phys. Chem. Chem. Phys.* **2004**, *6*, 1966–1969.
(11) Varma, S.; Rempe, S. B. *Biophys. Chem.* **2006**, *124*, 192–199.
(12) Ramaniah, L.; Bernasconi, M.; Parrinello, M. *J. Chem. Phys.* **1999**, *111*, 1587–1591.

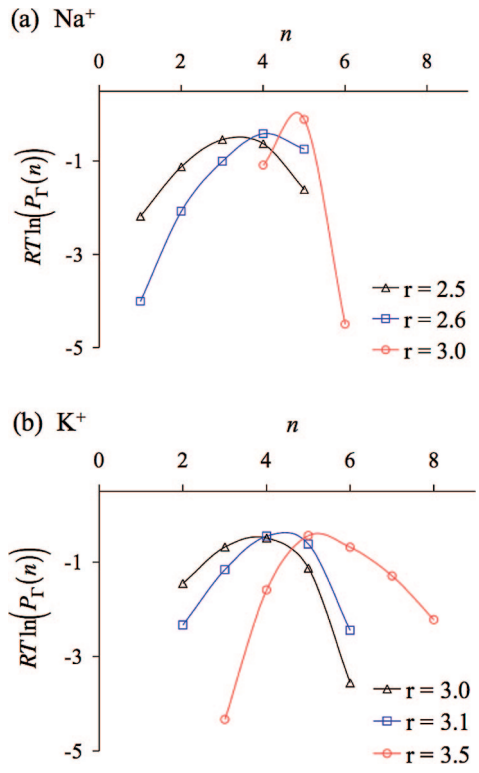


Figure 1. Probability $P_\Gamma(n)$ to find n water oxygens within a spherical subvolume $\Gamma = 4/3\pi r^3$ around the ion, where r is the distance from the ion in Å. These probabilities were estimated from separate AIMD trajectories of (a) Na^+ and, (b) K^+ ions. Probabilities are plotted in energy scale (kcal/mol).

finding greater than six water molecules within this subvolume becomes negligible (Figure 1). A similar trend is also seen for Na^+ ions, where decreasing the size of the subvolume by setting $r = 2.5$ Å reduces the preferred coordination number to three. In other words, there is a free energy cost ΔG_Γ associated with transferring ions from their respective preferred coordination states to a different coordination state inside a given Γ,

$$\Delta G_\Gamma = -RT \ln(P_\Gamma(n)/P_\Gamma(n_{\text{preferred}})) \quad (1)$$

This energetic cost gets progressively larger as n takes up values much larger or smaller than $n_{\text{preferred}}$.

There are, however, several exceptional cases in nature[1,2,13−19] where these ions readily partition from such aqueous-phase coordination states into spatial regions having much higher coordination numbers. For example, low-temperature X-ray data show that the binding sites in lysozymes, neurotransmitter transporters, dialkylglycine decarboxylases, and thrombins coordinate Na^+ ions using five or six ligands, and that all these ligands lie within 2.5 Å from the Na^+ ion (Table 1). At the same time, other biomolecules like potassium channels, nonactin, pyruvate kinases, and chaperones coordinate K^+ ions using more than six ligands, and all these ligands lie within 3.0 Å from the K^+ ion. In particular, crystallographic data for the celebrated potassium channels[20−23] show that their binding sites coordinate K^+ using eight carbonyl ligands, all of which are within 3.0 Å from the K^+ ion. This makes for twice as many ligands as seen preferentially around K^+ in aqueous phase, which according to eq 1 and the data in Figure 1, indicates that the transfer of a K^+ from liquid water to the binding sites in K-channels requires an enormous uphill transition on the free energy surface.

Table 1. Examples of Proteins That Bind Na^+ and K^+ Ions in States of Over-Coordination (Higher than the Coordination in Water)

biomolecule	PDB ID	no. ligands	r_{ave} (Å)
	Na^+		
neurotransmitter transporter	2A65[a]	5	2.3
dialkylglycine decarboxylase	2DKB[b]	6	2.4
lysozyme	1V7T[c]	6	2.4
neurotransmitter transporter	2A65[a]	6	2.4
thrombokinase	1BOK[d]	6	2.5
	K^+		
chaperone Hsc70	1HPM[e]	7	3.0
pyruvate kinase	1A49[f]	7	2.8
chaperone Hsc70	1HPM[e]	8	2.9
potassium channel	1K4C[g]	8	2.8
nonactin	n.a.[h]	8	2.8

[a] Reference 13. [b] Reference 14. [c] Reference 15. [d] Reference 16. [e] Reference 18. [f] Reference 17. [g] Reference 20. [h] Reference 19.

Clearly, there cannot be large free energy costs associated with such processes as ions are readily able to transfer back-and-forth between aqueous phase and biological molecules under standard physiological conditions. What physical principles explain such structural transitions? Do the binding site environments in these biological molecules somehow trigger a transition in ion coordination preferences? Or are there any special chemical and/or structural features present in the binding environments of biological molecules that simply reduce the energetic cost associated with overcoordination (relative to the low coordinations in water), without actually altering ion coordination preferences? Or do the specific chemistries of the ligands present in these binding sites effect such structural transitions?

An approach toward understanding the underlying basis of this ion coordination phenomenon begins with identifying the roles of various factors that affect the energetics of ion−ligand association reactions in solutions, such as ligand number, solvation environment, and ligand chemistry. This task becomes easier if the energetics of the ion−ligand association reactions in solutions could be estimated as functions of such determinants, thus providing insights into their explicit and combined roles and leading to clues for explaining the above anomaly. One way to accomplish this is to utilize the quasi-chemical

(13) Yamashita, A.; Singh, S. K.; Kawate, T.; Jin, Y.; Gouaux, E. *Nature* **2005**, *437*, 215–223.
(14) Toney, M. D.; Hohenester, E.; Cowan, S. W.; Jansonius, J. N. *Science* **1993**, *261*, 756–759.
(15) Harata, K.; Akiba, T. *Biol. Cryst.* **2004**, *60*, 630–637.
(16) Scharer, K.; Morgenthaler, M.; Paulini, R.; Obst-Sander, U.; Banner, D. W.; Schlatter, D.; Benz, J.; Stihle, M.; Diederich, F. *Angew. Chem., Int. Ed.* **2005**, *44*, 4400–4404.
(17) Larsen, T. M.; Benning, M. M.; Rayment, I.; Reed, G. H. *Biochemistry* **1998**, *37*, 6247–6255.
(18) Wilbanks, S. M.; McKay, D. B. *J. Biol. Chem.* **1995**, *270*, 2251–2257.
(19) Dobler, M. *Ionophores and Their Structures*; John Wiley & Sons, Inc.: New York, 1981; p 379.
(20) Zhou, Y.; Morais-Cabral, J. H.; Kaufman, A.; MacKinnon, R. *Nature* **2001**, *414*, 43–48.
(21) Jiang, Y.; Lee, A.; Chen, J.; Cadene, M.; Chait, B. T.; MacKinnon, R. *Nature* **2002**, *417*, 515–522.
(22) Jiang, Y.; Lee, A.; Chen, J.; Ruta, V.; Cadene, M.; Chait, B. T.; MacKinnon, R. *Nature* **2003**, *423*, 33–41.
(23) Kuo, A.; Gulbis, J. M.; Antcliff, J. F.; Rahman, T.; Lowe, E. D.; Zimmer, J.; Cuthbertson, J.; Ashcroft, F. M.; Ezaki, T.; Doyle, D. A. *Science* **2003**, *300*, 1922–1926.

organization of solution theory.[24−30] In this theory, the chemical potential of the solute is formulated in such a way that the 3-dimensional space around the solute is divided into separate inner- and outer-shell domains, enabling an estimation of energetic contributions to ion−ligand association arising from local and distant solvation environments.

Using the quasi-chemical approach, we first estimate solvation free energies of Na^+ and K^+ ions in three different liquid media, each representing a different ligand chemistry offered by the binding sites in biological molecules. Such a strategy provides estimates of the explicit roles of ligand chemistry and solvation phase in driving ion−ligand association reactions. In addition, the inner-shell clusters in the cluster implementation are built up by incrementally solvating the ion with increasing numbers of surrounding solvent molecules, which provides an estimate of the role of ligand numbers in driving ion solvation. We then carry out a series of other case studies using classical molecular dynamics (MD) simulations, probing the inferences derived from the earlier quasi-chemical analysis. Finally, we scrutinize our findings from the quasi-chemical and MD analyses against the structural and chemical properties of the ion-binding sites listed in Table 1, where ion binding occurs in states of overcoordination relative to water.

We find that the local solvation environment of ligands, which also includes ligand concentrations and structural flexibility, plays a significant role in driving transitions in ion coordination structures. Although important to other aspects of ion−host interactions, ligand chemistry appears to be less important in driving ion coordination transitions.

2. Methods

2.1. Quasi-Chemical Theory of Solutions. The quasi-chemical organization of solution theory[24−30] enables calculation of solvation free energy in terms of individual contributions arising from local and distant solvation effects. The volume surrounding the solute is divided into an inner- and an outer-shell domain, where the inner-shell domain captures the solute and a subset of the solvent molecules. Solvent occupation statistics are then collected for the inner domain, where clustering equilibria occur, thus providing an estimate of the effect of coordination number on solvation energy. Since details of the theory are described elsewhere,[24−30] here we outline only those results of the formulation that show how the free energy is partitioned into individual components, which provide insight into the driving forces of ion solvation by any medium.

The Gibbs free energy or the chemical potential of a solute particle (A) can be written as,[30,31]

$$\mu_A = RT \ln\left(\frac{\rho_A \Lambda_A^3}{q_A^{int}}\right) + \mu_A^{ex} \tag{2}$$

where ρ_A is the number density of the solute, Λ_A is the thermal de Broglie wavelength and q_A^{int} is the partition function of a single solute particle without translational degrees of freedom. The first term in eq 2 is the ideal gas contribution to the chemical potential, while the second term, which is the quantity of primary interest for the present work, is the excess chemical potential. The potential distribution theorem[25,30] defines this excess chemical potential as

$$\mu_A^{ex} = -RT \ln\langle\langle e^{-\beta \Delta U_A}\rangle\rangle_0 \tag{3}$$

where, $\beta = 1/RT$ and ΔU_A is the potential energy of interaction between the solute and solvent molecules. The double brackets indicate averaging over the thermal motions of the solute and the solvent molecules. The subscript zero, in line with Widom's particle insertion method,[25] emphasizes that this average is obtained in the absence of solute−solvent interactions. This excess chemical potential can be separated into two terms:[24−27,29,30]

$$\beta\mu_A^{ex} = \ln[x_{\Gamma,n=0}] - \ln\left\langle\left\langle e^{-\beta \Delta U_A} \prod_j [1 - b_\Gamma(j)]\right\rangle\right\rangle_0 \tag{4}$$

The first term provides the inner-shell contribution to the excess chemical potential of solute A, while the second term provides the outer-shell contribution. The quantity $x_{\Gamma,n=0}$ in the first term is the probability that an arbitrarily defined inner-shell domain Γ contains the solute particle without any solvent molecules ($n = 0$). $b_\Gamma(j)$ is an indicator function that takes up a value of one when the jth solvent molecule is inside Γ and zero otherwise. This formulation can be utilized directly in conjunction with standard molecular simulation techniques, such as MD or Monte Carlo simulations, to determine solvation energies as functions of, for example, the shapes and volumes[32] of the inner-shell domain Γ. It is straightforward to see that while the net excess chemical potential of a solute μ_A^{ex} will be independent of the shape or the volume of the inner-shell domain Γ, as can also be seen in Figure 3 of reference 32, the relative contributions arising from inner- and outer-shell domains will change with the definition of Γ.

For the present study, we are particularly interested in probing solvation free energies as functions of the numbers ($n > 0$) of solvent molecules inside the inner-shell domain of the solute. This can be accomplished by expressing the probability $x_{\Gamma,n=0}$ as a function of the equilibrium constants K_n of the solute−solvent association reactions that occur inside the inner-shell domain of the solute particle. In a specific scenario where solute A interacts with n solvent particles of type B inside volume Γ to form cluster AB_n,

$$A + nB \xrightarrow{K_n} AB_n \tag{5}$$

the probability to find zero solvent molecules within Γ is,

$$x_{\Gamma,n=0} \equiv \frac{C_{AB_0}}{\sum_n C_{AB_n}} = \frac{1}{\sum_n K_n [C_B]^n} \tag{6}$$

Here C_{AB_n} and C_B are, respectively, the concentrations of the cluster and solvent particles. Inserting this into eq 4 yields

$$\beta\mu_A^{ex} = -\ln\left[\sum_{n\geq 0} K_n C_B^n\right] - \ln\left\langle\left\langle e^{-\beta \Delta U_A} \prod_j [1 - b_\Gamma(j)]\right\rangle\right\rangle_0 \tag{7}$$

For the specific association reactions defined by eq 5, the equilibrium constants K_n can be expressed as[30]

$$K_n = K_{n,\Gamma}^{(0)} \frac{\langle\langle e^{-\beta \Delta U_{AB_n}}\rangle\rangle_{0,\Gamma}}{\langle\langle e^{-\beta \Delta U_{AB_0}}\rangle\rangle_{0,\Gamma} \{\langle\langle e^{-\beta \Delta U_B}\rangle\rangle_0\}^n} \tag{8}$$

(24) Guggenheim, E. A. *Proc. Royal Soc. London., Ser. A* **1935**, *148*, 304–312.
(25) Widom, B. *J. Phys. Chem.* **1982**, *86*, 869–872.
(26) Pratt, L. R.; LaViolette, R. A. *Mol. Phys.* **1998**, *95*, 909–915.
(27) Martin, R. L.; Hay, P. J.; Pratt, L. R. *J. Phys. Chem. A* **1998**, *102*, 3565–3573.
(28) Hummer, G.; Pratt, L. R.; Garcia, A. E. *J. Phys. Chem. A* **1998**, *102*, 7885–7895.
(29) Pratt, L. R.; Rempe, S. B. Simulation and Theory of Electrostatic Interactions in Solution. In *Quasi-Chemical Theory and Implicit Solvent Models for Simulations*; Pratt, L. R., Hummer, G., Eds.; AIP Conference Proceedings, 1999; pp 172−201.
(30) Beck, T. L.; Paulaitis, M. E.; Pratt, L. R. *The Potential Distribution Theorem and Models of Molecular Solutions*; Cambridge University Press: New York, 2006; p 244.
(31) Hansen, H. P.; McDonald, I. R. *Theory of Simple Liquids*; Academic Press: New York, 1986.
(32) Sabo, D.; Varma, S.; Martin, M.; Rempe, S. B. *J. Phys. Chem. B* **2008**, *112*, 867–876.

Here, $K^{(0)}_{n,\Gamma}$ are the "ideal" equilibrium constants of the association reactions given by eq 5. They differ from K_n as they do not include any effects of the complementary region Γ^C outside the inner-shell domain. The specific effects of Γ^C on K_n are incorporated through the ensemble averages in eq 8. Note that the ideal equilibrium constant $K^{(0)}_{n,\Gamma}$ should not be confused with the gas-phase equilibrium constant $K^{(g)}_n$. $K^{(0)}_{n,\Gamma}$ are obtained under the condition that the association equilibrium occurs within a predefined inner-shell domain Γ, whereas the estimations of $K^{(g)}_n$ are free from any such impositions. The quantity $\langle\langle e^{-\beta\Delta U_{AB_n}}\rangle\rangle_{0,\Gamma}$ is the ensemble average of the distribution that is the product of the distributions for the solvent molecules in Γ^C and for the complex enclosed within the region Γ. The ensemble average $\langle\langle e^{-\beta\Delta U_B}\rangle\rangle_0$, just as in eq 3, is related to the excess chemical potential of the solvent molecule B as:

$$\mu^{ex}_B = -RT\ln\langle\langle e^{-\beta\Delta U_B}\rangle\rangle_0 \tag{9}$$

The ensemble average $\langle\langle e^{-\beta\Delta U_{AB_0}}\rangle\rangle_{0,\Gamma}$ in eq 8 is, in fact, precisely equal to the ensemble average in the second term of eq 7, that is,

$$\langle\langle e^{-\beta\Delta U_{AB_0}}\rangle\rangle_{0,\Gamma} = \left\langle\left\langle e^{-\beta\Delta U_A}\prod_j [1 - b_\Gamma(j)]\right\rangle\right\rangle_0 \tag{10}$$

Introducing the relationships in eqs 8 and 10 into eq 7 and rearranging yields

$$\beta\mu^{ex}_A = -\ln\left[\sum_{n\geq 0} K^{(0)}_{n,\Gamma}\left[\frac{\langle\langle e^{-\beta\Delta U_{AB_n}}\rangle\rangle_{0,\Gamma}}{\{\langle\langle e^{-\beta\Delta U_B}\rangle\rangle_0\}^n}\right]C^n_B\right] \tag{11}$$

This quasi-chemical formulation shows how the excess chemical potential of a solute A can be estimated in terms of individual contributions arising from inner- and outer-shell domains: $K^{(0)}_{n,\Gamma}$ provides the contribution from the inner-shell domain, while the ratio of the ensemble averages within the inner square brackets provides the contribution from the outer-shell domain.

We now describe how we implement this quasi-chemical framework to determine the solvation free energies of Na^+ and K^+ ions in three different liquid media: methanol, water and formamide. Under standard conditions, these solvents have different dielectric constants of 33, 78.5, and 109.5, respectively. In addition, they provide different chemistries for ion−ligand association that are representative of the chemistries offered to ions in the binding sites of biological molecules.

2.2. Implementation. Following the formulation in eq 11, the solvation free energy, $\Delta G^A_n(l_B)$, of an ion of type A in the liquid phase (l) of solvent molecules of type B can be computed as the sum of four separate terms:[24−27,29,30]

$$\Delta G^A_n(l_B) = \Delta G^A_{n,\Gamma}(g_B) + n\Delta G^B_{conc} + \Delta G^{AB_n}_{n,\Gamma}(l_B) - n\Delta G^B(l_B) \tag{12}$$

These four terms are:

(a) The "intrinsic" free energies $\Delta G^A_{n,\Gamma}(g_B)$ of ion−ligand association reactions estimated in the gas phase, but under the condition that the ion−ligand association equilibria occur within a predefined volume Γ around the ion. $\Delta G^A_{n,\Gamma}(g_B)$ can occasionally be equal to the ion−ligand association energies in gas phase, but in general $\Delta G^A_{n,\Gamma}(g_B) \geq \Delta G^A_n(g_B)$.

(b) The contribution from the actual concentration of the solvent molecules, ΔG^B_{conc}, which may be different from their concentrations in gas phase used in the estimation of the previous term $\Delta G^A_{n,\Gamma}(g_B)$.

(c) The solvation free energies, $\Delta G^{AB_n}_{n,\Gamma}(l_B)$, or the excess chemical potentials of individual clusters (AB_n) in the solvent l_B.

(d) The solvation free energy, $\Delta G^B(l_B)$, or the excess chemical potential of the solvent molecule B in its bath l_B.

Note that while we evaluate solvation free energies, $\Delta G^A_n(l_B)$, for various ion−ligand coordination complexes, an ion's excess chemical potential will result from formation of the most probable complex, that is, the one for which $\Delta G^A_{n=n_{preferred}}(l_B)$ has the lowest most favorable value.

One way to determine these contributions to ion solvation is to utilize the *ab initio* level of theory to describe explicitly the complex interactions between the ion and the ligands inside the inner-shell domain Γ, while the remaining longer-range interactions can be described using an implicit solvent model. The radius chosen to define the spherical boundary for the inner-shell domain essentially determines the number of solvent molecules that will be treated explicitly at the *ab initio* level of theory. Choosing a large radius can result in expensive calculations on a large number of association reactions, whereas choosing a small radius may not provide sufficient information to understand the dependency of solvation energy on relevant ligand numbers. Note that while the net solvation free energy of an ion, $\Delta G^A_n(l_B)$, will not depend upon the choice of the radius of the inner-shell domain Γ, the relative contributions arising from the inner- and outer-shell domains will change with the choice of the radius.

To select a radius for defining the spherical boundary of the inner-shell domain, we first examine the low temperature X-ray data in Table 1. We see that, despite the apparent overcoordination in the biological ion-binding sites, the average distance of ligands from Na^+ is less than 2.6 Å, while the average distance of ligands from K^+ is less than 3.1 Å. From the AIMD simulations in liquid water,[8−11] we also know that the average distances of Na^+ and K^+ ions from the water oxygens within their respective canonical inner-shells are also less than 2.6 Å and 3.1 Å. We therefore define the inner-shell boundary around Na^+ using a radius of 2.6 Å, and around K^+ using a radius of 3.1 Å. As we will demonstrate using molecular dynamics simulations, these boundaries also naturally accommodate the average inner-shell ion−carbonyl distances for fully flexible ligands under conditions of room temperature. Note that for these chosen volumes of the inner-shell domains, the packing penalties are small[32] and expected to be equivalent in magnitude but opposite in sign to the van der Waals interaction between the inner- and outer-shell molecules. Neither of these contributions is determined explicitly in this work as they are expected to cancel out.

The intrinsic Gibbs free energies $\Delta G^A_{n,\Gamma}(g_B)$ of the ion−ligand association reactions are determined by describing all interactions at the *ab initio* level of theory. The three-dimensional structures of the ligands and their respective inner-shell coordination complexes with the ions were first optimized using the hybrid B3LYP[33,34] density functional theory (DFT). All ion−ligand coordination distances were found to be less than the respective radii chosen to describe the boundary of the inner-shell domain. See supplementary text for illustration of the Na^+−water and K^+−water optimized configurations and note that the configurations of the ion−methanol and ion−formamide optimized complexes resemble those of the ion−water optimized configurations.

Each optimized structure is then analyzed to determine its individual free energies $G^A(g)$, $G^B(g)$, and $G^{AB_n}_{n,\Gamma}(g)$. Entropies for all reactant and product species were determined separately from partition functions[35] evaluated at their respective potential energy minima and at a temperature (T) of 298 K and pressure (P) of 1 atm. In particular, harmonic expansions (Hessian analysis) about these potential energy minima produce estimates of the vibrational frequencies, thereby capturing the effects of thermal fluctuations within each potential energy well. If optimization algorithms fail and the final geometry does not correspond to a stable minimum on the potential energy surface, the computed vibrational frequencies end up imaginary. All frequencies were found to be real, confirming the existence of energetically optimized structures. Next, these free energies were combined to obtain gas-phase reaction free energies for the various inner-shell association reactions,

$$\Delta G^A_{n,\Gamma}(g_B) = \Delta G^{AB_n}_{n,\Gamma}(g) - G^A(g) - nG^B(g) \tag{13}$$

(33) Becke, A. D., III *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(34) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
(35) McQuarrie, D. A. *Statistical Mechanics*, 2nd ed.; University Science Books: Mill Valley, CA, 2000; p 641.

Table 2. Effect of Na^+ and K^+ Basis Sets on Computed Gas Phase Gibbs Free Energies (kcal/mol)[a]

	$\Delta G^{Na^+(H_2O)_3 \rightarrow Na^+(H_2O)_2+(H_2O)}(g)$		$\Delta G^{K^+(H_2O)_3 \rightarrow K^+(H_2O)_2+(H_2O)}(g)$	
basis set	estimated	% error	estimated	% error
expt.[b]	9.5		6.3	
6-31G*	9.6	1.1	5.8	7.9
6-31+G(d)	9.6	1.1	5.6	11.1
6-31+G*	−	−	5.4	14.3
6-311+G*	−	−	6.4	1.6

[a] Our goal here is to find the smallest basis set that estimates the experimental free energy difference between two randomly selected ion−water complexes within a 2% error. To accomplish this, we start with the 6-31G* basis set for each ion and gradually increase its size. We keep the basis set for the water molecule fixed at 6-31++G** in accord with the observation[37] that it yields the energetic and structural properties of dipole−dipole interactions, H_2O-H_2O and H_2O-HF, in excellent agreement with experiment. [b] Reference 39.

Note that in these computations we do not account for: (a) anharmonicities in the vibrational motions of the ion complexes and (b) fluctuations due to configurational changes within the inner-shell domain, as each of the different *n*-fold inner-shell coordination states of the ions are represented using single optimized structures. Separate calculations in fact demonstrate a significant effect on energetics due to the presence of anharmonicity and configurational fluctuations only in higher-order ($n > 5$) ion−ligand clusters. While corrections for these anharmonicities and configurational fluctuations result in more exothermic association reactions, they do not alter any of the conclusions drawn from this study. The details of these calculations are not included here as they are subject of a separate publication. (See reference 32 for a description of methods used for explicitly calculating free energy contributions due to anharmonicities.)

All optimization and subsequent frequency calculations were carried out with the *Gausssian 03* suite of programs[36] using the following basis sets: a 6-31G* basis set for Na^+ ions, a 6-31++G** basis set for H, C, N, and O atoms, and a 6-311+G* basis set for K^+ ions. Our reasons for selecting these basis sets are summarized in the legend of Table 2. Guided by previous investigations dealing with the effects of basis sets on estimating gas-phase H_2O-H_2O and H_2O-HF interactions,[37] we systematically explored several possibilities until we arrived at this combination of basis sets, which yields ion−water dissociation energies in excellent agreement with experiment.[38−40]

The favorable gain in reaction free energy ΔG^{B}_{conc} associated with a higher concentration (or, alternatively, higher pressure) of solvent molecules in the liquid phase ($C_l > C_g$) was determined as follows. Since the translational term of the molecular partition function can be singled out,[30,35,41] allowing the chemical potential of a solvent molecule B to be written as a summation,

$$\mu_B(l) = \mu_B(g) - RT \ln\langle e^{-\Delta U_B/kT}\rangle_N - RT \ln(C_l/C_g) \qquad (14)$$

this translational contribution is determined as a function of the ratio of solvent molecule concentrations (or pressures) in liquid (C_l) and gas (C_g) phases, that is, as

$$\Delta G^{B}_{conc} = -\beta^{-1} \ln(C_l/C_g) \qquad (15)$$

The gas phase concentration was taken as 0.041 M, while the liquid phase concentrations of methanol, water, and formamide were taken as 24.7, 55.6, and 25.2 M, respectively.

The excess chemical potentials of the solvent molecules, $\Delta G^{B}(l_B)$, and individual clusters, $\Delta G^{AB_n}_{n,\Gamma}(l_B)$, are determined by treating the outer-shell domain as an implicit solvent. The finite difference (FD) scheme of the APBS package[42] with a finest grid spacing of 0.1 Å was used for solving the Poisson's equation. Partial charges of atoms in the ligands and complexes were obtained from separate DFT/B3LYP calculations using the optimized structures and the ChelpG method,[43] and were distributed over the FD grids using cubic B-spline discretization. Atomic radii required for defining solvent exclusion regions were taken from Stefanovich and Truong;[44] and the diameters of the solvent molecules required to create the molecular surfaces on the FD grids were taken as the distances corresponding to the principal maxima of their oxygen−oxygen pair distribution functions in liquid state, that is, 1.4 Å for water, 1.4 Å for methanol[45] and 2.15 Å for formamide.[46] The radii needed to partition the inner- and outer-shell domains were taken as 2.6 Å for Na^+ ions and 3.1 Å for K^+ ions.

We note that the method used above for estimating atomic partial charges does not account for the perturbation of the solvent reaction field onto the solute Hamiltonian. To test the effect of the solvent reaction field on the estimation of partial charges and ultimately on the estimation of the excess chemical potentials of clusters, $\Delta G^{AB_n}_{n,\Gamma}(l_B)$, we utilized the integral equation formalism polarizable continuum model[47] (IEFPCM) implemented in the *Gausssian 03* suite of programs[36] and computed the excess chemical potentials for two test cases, $\Delta G^{Na^+W_4}_{4,\Gamma}(l_W)$ and $\Delta G^{K^+W_4}_{4,\Gamma}(l_W)$. Consistent with other studies of closed-shell systems,[48] we find that the incorporation of the perturbation of the solvent reaction field onto the solute Hamiltonian has a minimal effect on the results.

We also note that treating the outer-shell domain Γ^C as an implicit solvent can lead to errors in the estimation of free energy components $\Delta G^{AB_n}_{n,\Gamma}(l_B)$ and $\Delta G^{B}(l_B)$.[49−51] The broad phenomena we observe in this work, however, appear sufficiently robust and are not likely to be compromised by this inaccuracy. In addition, the treatment of the outer-shell domain as an implicit solvent is expected to result in the overestimation of both terms, resulting in a partial-to-complete cancellation when the final solvation free energies $\Delta G^{A}_{n}(l_B)$ are computed. For example, the calculation of $\Delta G^{B}(l_B)$ in the case of ion solvation in water yields a value of −8.4 kcal/mol, which in comparison to other values,[52] is overestimated by ~2 kcal/mol. The radii chosen to define the dielectric boundaries, that is, 2.6 Å for Na^+ ions and 3.1 Å for K^+ ions, places the low dielectric boundary of the cluster in such a way that it protrudes into the high dielectric region, causing the solvent in the outer-shell domain to be physically closer to the cluster than seen in AIMD simulations.[8−11] This should lead to an overestimation of the values of $\Delta G^{AB_n}_{n,\Gamma}(l_B)$, and as the cluster size increases, this overestimation should also increase, ultimately resulting in a minimization of errors when the summation $\Delta G^{AB_n}_{n,\Gamma}(l) - n\Delta G^{B}(l)$ is computed in eq 11.

(36) Frisch, M. J.; *Gaussian 03*; Gaussian, Inc.: Wallingford CT, 2004.
(37) Salvador, P.; Paizs, B.; Duran, M.; Suhai, S. *J. Comput. Chem.* **2001**, *22*, 765–786.
(38) Dzidic, I.; Kebarle, P. *J. Phys. Chem.* **1970**, *74*, 1466–1474.
(39) Blades, A. T.; Klassen, J. S.; Kebarle, P. *J. Am. Chem. Soc.* **1996**, *118*, 12437–12442.
(40) Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R. J. *J Phys. Chem. A* **1998**, *102*, 7787–7794.
(41) Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 792–803.
(42) Baker, A. N.; Sept, D.; Joseph, S.; Holst, M.; McCammon, A. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037–10041.
(43) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361–373.
(44) Stefanovich, E. V.; Truong, T. N. *Chem. Phys. Lett.* **1995**, *244*, 65–74.
(45) Yamaguchi, T.; Hidaka, K.; Soper, A. K. *Mol. Phys.* **1999**, *96*, 1159–1168.
(46) Tsuchida, E. *J. Chem. Phys.* **2004**, *121*, 4740–4746.
(47) Cances, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 1032–1041.
(48) Marenich, A. V.; Cramer, C. J.; Truhlar, D. J. *J. Chem. Theory Comput.* **2008**, *4*, 877–887.
(49) Simonson, T. *Rep. Prog. Phys.* **2003**, *66*, 737–787.
(50) Baker, A. N. *Curr. Opin. Struct. Biol.* **2005**, *15*, 137–143.
(51) Varma, S.; Jakobsson, E. *Biophys. J.* **2004**, *86*, 690–704.
(52) Asthagiri, D.; Pratt, L. R.; Ashbaugh, H. S. *J. Chem. Phys.* **2003**, *119*, 2702–2708.

2.3. *Ab Initio* Molecular Dynamics. Two separate AIMD trajectories were generated within the Born−Oppenheimer scheme[53] using the PW91 functional[54] implemented in VASP[55] version 4.2: one involving a Na^+ ion embedded in a box of 64 waters, while the other involving a K^+ ion embedded in a box of 64 waters. Both trajectories are extensions of previously reported trajectories[11] to longer time scales of 20 ps for Na^+ and 40 ps for K^+. Note that the K^+−O pair distribution functions from the 40 ps trajectory of K^+ have been reported elsewhere.[56] The following parameters were used for running the simulations in NVE ensembles and with periodic boundaries: cubic box of length 12.4171 Å to emulate water density under standard conditions; Ewald summation with background neutralizing charge; projector augmented-wave method[57,58] for description of core-valence electron interactions; explicit treatment of the 2P electrons of Na^+ and the 3P electrons of K^+; deuterium mass assigned to protons; kinetic energy cutoff of 500 eV for the expansion of plane waves; gamma-point Brillouin zone sampling; cutoff of 10^{-6} eV for energy convergence; and an integration time step of 0.5 fs. The average temperature in the simulation with Na^+ was 294 ± 11 K, whereas it was 313 ± 21 K for the simulation with K^+. Data from these trajectories were used for estimating the probabilities illustrated in Figure 1.

2.4. Classical Molecular Dynamics. All classical molecular dynamics simulations were carried out using GROMACS[59] version 3.1.4. Unless otherwise stated, the following parameters were used to carry out simulations in an NVT ensemble: cubic box of length 30 Å; periodic boundaries in all directions; PME with real space cutoff of 10 Å, background neutralizing charge, dipole correction, Fourier spacing of 1.5 Å, a sixth order interpolation and a tolerance of 10^{-5} units; a cutoff of 10 Å for van der Waals interactions; recently revised van der Waals parameters[60] for K^+; SPC/E[61] charges and van der Waals parameters for description of water models; all atom OPLS[62] charges and van der Waals parameters for description of all other types of atoms and molecules; an integration time-step of 2 fs; the SETTLE algorithm[63] for constraining bond lengths in water molecules; the LINCS algorithm[64] to constrain all other bond lengths; and a Berendsen method[65] with a coupling constant of 0.1 ps to maintain temperature at 298.15 K. In simulations involving only an ion and water molecules, the box was filled with 893 water molecules to emulate the density of water under standard conditions. When other molecules were introduced in this simulation box, appropriate numbers of water molecules were deleted to maintain the water densities.

3. Results and Discussion

3.1. Quasi-Chemical Analysis: Determinants of Ion Solvation. To determine how ions are able to partition from liquid water (l_W) to higher coordinations in the binding sites of some proteins (ξ) without uphill transitions on the free energy surface, we first need to understand the explicit roles of the various factors that drive the energetics of ion−ligand association reactions in solution. To this end, we first utilize the quasi-chemical framework of solution theory[24−30] to estimate the solvation free energies of Na^+ and K^+ ions in liquid water. As described in the methods section, the solvation free energies are estimated as functions of coordination numbers and in terms of individual contributions arising from local and distant solvation effects, which provide an opportunity to inspect their individual and combined roles during ion solvation.

Table 3. Quasi-Chemical Components (in kcal/mol) in the Calculation of Hydration Free Energy of Na^+ and K^+ Ions

A	n	$\Delta G^{A}_{n,\Gamma}(g_W)$[a] [Exp.][b]	$n\Delta G^{W}_{conc}$[c]	$\Delta G^{A}_{n,\Gamma}(g_W) + n\Delta G^{W}_{conc}$	$\Delta G^{AW_n}_{n,\Gamma}(l_W)$[d]	$n\cdot\Delta GW(l_W)$[e]	$\Delta G^{A}_{n}(l_W)$[f]
Na^+	1	−18.6 [−18.8]	−4.3	−22.8	−61.2	−8.4	−75.6
	2	−34.2 [−31.9]	−8.5	−42.7	−60.2	−16.7	−86.2
	3	−43.3 [−40.9]	−12.8	−56.1	−59.8	−25.1	−90.9
	4	−49.4 [−46.9]	−17.1	−66.5	−60.2	−33.5	−93.3
	5	−49.0 [−]	−21.4	−70.4	−59.6	−41.9	−88.1
	6	−45.6 [−]	−25.6	−71.3	−59.1	−50.2	−80.1
K^+	1	−12.0 [−11.8]	−4.3	−16.3	−52.9	−8.4	−60.8
	2	−21.7 [−20.7]	−8.5	−30.3	−53.7	−16.7	−67.2
	3	−27.7 [−27.0]	−12.8	−40.6	−55.0	−25.1	−70.4
	4	−32.0 [−]	−17.1	−49.0	−56.3	−33.5	−71.8
	5	−29.8 [−]	−21.4	−51.1	−56.1	−41.9	−65.4
	6	−26.9 [−]	−25.6	−52.5	−54.3	−50.2	−56.6
	7	−28.8 [−]	−29.9	−58.7	−54.9	−58.6	−55.0
	8	−30.6 [−]	−34.2	−64.8	−56.2	−67.0	−53.9

[a] $\Delta G^{A}_{n,\Gamma}(g_W)$ is the intrinsic Gibbs free energy of ion−ligand association reaction at T = 298 K and P = 1 atm, and is estimated using eq 13. Computed values are compared to the experimental estimates of Tissandier et al.,[40] but the comparison is limited to lower-order clusters, $n \le 4$ for Na^+ and $n \le 3$ for K^+. This is because separate calculations[67] at the MP2 level of theory show that the experimental estimates for higher-order gas-phase clusters correspond to configurations where waters do not directly coordinate the ion, a scenario entirely different from the one considered for this quasi-chemical analysis. [b] Reference 40. [c] ΔG^{W}_{conc} refers to the favorable gain in reaction free energy associated with a higher concentration of water molecules in the liquid phase as compared to the gas phase. [d] $\Delta G^{AW_n}_{n,\Gamma}(l_W)$ refers to the solvation free energy of an n-fold ion−water cluster in liquid water. [e] $\Delta G^{W}(l_W)$ is the solvation free energy of a single water molecule W in liquid water. [f] $\Delta G^{A}_{n}(l_W)$ is the solvation free energy of the ion coordinating with n water molecules in liquid water and is equal to the summation given by eq 12. The lowest most favorable value of $\Delta G^{A}_{n}(l_W)$ corresponds to the ion's absolute hydration free energy as it is the dominant term.

3.1.1. Ion Solvation in Liquid Water. The results of the calculation of the solvation free energies of Na^+ and K^+ ion in liquid water using the quasi-chemical framework of solution theory[24−30] are given in Table 3.

For the specific radii used to define the inner-shell domains, that is, 2.6 Å for Na^+ and 3.1 Å for K^+, we find that both ions prefer to coordinate with four water molecules, consistent with the results from AIMD simulations illustrated in Figure 1. (Recall that these preferential coordinations are distinct from average coordination values defined by the principal minima of radial distribution functions.) The resulting absolute hydration free energies of both ions, that is, values corresponding to $\Delta G^{A}_{n}(l_W)$ when $n = n_{preferred} = 4$, are within 3% of their respective experimental estimates.[66]

For Na^+, the computed absolute hydration free energy $\Delta G^{Na^+}_{4}(l_W) = -93.3$ kcal/mol is only slightly overestimated with respect to the experimental value[66] of −90.8 kcal/mol. This is perhaps due in part to the slightly larger[38,40] value of the predicted gas phase Na^+-water association free energy

(53) Born, M.; Oppenheimer, J. R. *Ann. Phys.* **1927**, *84*, 457–484.
(54) Perdew, J. P. *Unified Theory of Exchange and Correlation Beyond the Local Density Approximation*; Akademie Verlag: Berlin, 1991.
(55) Kresse, G.; Hafner, J. *Phys. Rev. B* **1993**, *47*, 558–561.
(56) Varma, S.; Rempe, S. B. *Biophys. J.* **2007**, *93*, 1093–1099.
(57) Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953–17979.
(58) Kress, G.; Joubert, D. *Phys. Rev. B* **1998**, *59*, 1758–1775.
(59) Lindahl, E.; Hess, B.; Spoel, D. *J Mol. Mod.* **2001**, *7*, 306–313.
(60) Varma, S.; Chiu, S. W.; Jakobsson, E. *Biophys. J.* **2006**, *90*, 112–123.
(61) Berendsen, H. J.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
(62) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
(63) Miyamoto, S.; Kollman, P. *J. Comput. Chem.* **1992**, *13*, 952–962.
(64) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
(65) Berendsen, H. J.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(66) Marcus, Y. *Biophys. Chem.* **1994**, *51*, 111–127.
(67) Feller, D.; Glendening, E. D.; Woon, D. E.; Feyereisen, M. W. *J. Chem. Phys.* **1995**, *103*, 3526–3542.

$\Delta G_{4,\Gamma}^{Na^+}(g_W) = \Delta G_4^{Na^+}(g_W)$ (see column 3 of Table 3) and in part to the use of implicit solvent in the treatment of long-range interactions.[32,52] These errors could be reduced[67] by switching to a Møller−Plesset[68] second-order perturbation theory (MP2) and/or by using explicit solvent molecules in the outer-shell domain Γ^C.[32]

The computed absolute hydration free energy of K^+, $\Delta G_4^{K^+}(l_W) = -71.8$ kcal/mol, is slightly underestimated with respect to the experimental value[66] of −74.1 kcal/mol. Unlike the case of Na^+, there is no experimental estimate for $\Delta G_{4,\Gamma}^{K^+}(g_W)$, but given that the values of $\Delta G_1^{K^+}(g_W)$, $\Delta G_2^{K^+}(g_W)$ and $\Delta G_3^{K^+}(g_W)$ are in excellent agreement with experiment,[38,40] it seems unlikely that this underestimation could result from a choice of functional or basis set. Treating solvent molecules in the outer domain explicitly may improve this value. Another option would be to extend the boundary to larger distances that include more solvent explicitly in the inner-shell domain, but then calculations become more time consuming, and the larger clusters that result may require explicit treatment of anharmonicity and configurational fluctuations. Nevertheless, correcting these small errors will not alter the trend in the dependence of solvation free energies $\Delta G_n^A(l_W)$ on coordination number observed for these specific inner-shell volumes, and is not the primary objective of this study.

We now assess the effect of the various terms that contribute to the solvation properties of these two ions. From the data in column 4 of Table 3, we see that ligand concentration plays a vital role in stabilizing ion−ligand complexes. When ligand concentration is increased from its value in gas phase (C_g = 0.041 M) to its value in liquid phase (C_W = 55.6 M), the stability of each gas phase ion−water cluster is enhanced, that is, the sum $\Delta G_{n,\Gamma}^A(g_W) + n\Delta G_{conc}^W$ in column 5 of Table 3 is less than the corresponding value of $\Delta G_{n,\Gamma}^A(g_W)$ in column 3. How is this relevant to ion binding in biological molecules?

Although biological molecules cannot or do not have the isotropic high densities of ligands found in liquids or solids, they can certainly ensure the presence of high local densities in their ion-binding regions through specific three-dimensional folds or topologies and increase their propensities to bind ions. Note that these densities do not correspond to the densities of ligands in the ion-complexed states but to the densities of ligands in the apo states of the binding sites. The need for high local densities to stabilize ion binding can therefore, at times, result in close similarities between the ion-complexed and apo-state structures of ion-binding sites. This aspect of ion binding has been observed for numerous proteins, with more recent examples coming from K-channels,[69] thrombins,[70] and cytosolic phospholipases,[71,72] where the structures of the apo and ion-complexed states are similar in the sense that they do not have different orders of ligand densities in their binding sites. Note also that the dependence of free energies on ligand concentrations is logarithmic, which allows binding sites the freedom to explore configurational space before changes in ligand densities begin to alter the energetics of ion binding appreciably. This freedom to explore configurational changes can be vital to biological molecules, especially to acclimatize to the presence or absence of high charges in their binding sites.[72]

From column 6 of Table 3, we see that the excess chemical potential $\Delta G_{n,\Gamma}^{AW_n}(l_W)$ of each ion−water cluster is favorable. In addition to the gas phase and the ligand concentration free energy terms in eq 12, that is, $\Delta G_{n,\Gamma}^A(g_W)$ and $n\Delta G_{conc}^W$, this term also serves to stabilize the formation of ion−water complexes in liquid water. Unlike the concentration term $\Delta G_{n,\Gamma}^{AW_n}(l_W)$, however, solvation of the clusters does not depend strongly on the coordination number of ions.

In the final step in the estimation of ion solvation free energies in liquid water, the excess chemical potential of the water molecule, $\Delta G^W(l_W)$, is multiplied by the total number of solvent molecules in the inner-shell domain (column 7 of Table 3) and subtracted out. This makes each successively larger coordination complex in liquid phase less favorable. More importantly, this solvent contribution disrupts the stabilities of larger clusters more than the smaller clusters, shifting the coordination preferences in favor of lower values. This process of subtracting out the excess chemical potential of a water molecule $\Delta G^W(l_W)$ from the total solvation free energy of an ion is physically analogous to accounting for the penalty $\Delta\Phi^W(l_W)$ associated with desolvating (extracting) a water molecule ligand from its aqueous environment to place it into a coordination complex, where $\Delta\Phi^W(l_W) = -\Delta G^W(l_W)$. There is a higher *overall* penalty associated with extracting larger numbers of water molecules from their natural bulk environment to place them in the coordination complex, which leads to a shift in the ion coordination preferences in favor of smaller numbers.

Note from the data in Table 3 that the structural and thermodynamic properties of ion solvation in liquid water arise from a coupling or a balance between contributions arising from both the inner- and outer-shell domains. If this coupling or balance is disturbed, such as due to changes in temperature and pressure and/or due to changes in properties of the local or distant environments, ion solvation properties may change. For example, it is well-known that increasing the temperature and pressure of salt solutions to very high values reduces the probability of the formation of higher-order coordinations to such an extent that the coordination preferences of ions shift in favor of lower values as compared to those found under ambient conditions of temperature and pressure.[73−77] The physical principles that drive these structural transitions, however, still remain unclear. In another well-known example, increasing the concentration of ions in water also results in reducing their average coordination numbers.[5,7,78] This result can, in fact, be understood in the context of the free energy terms given in eq 12. We shall take up this discussion in the next section, where we also argue that this observation is actually part of a broader phenomenon known as the Hofmeister effect.[79−82] Nevertheless, neither these nor any of the previous studies explain how Na^+

(68) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.
(69) Zhou, Y.; MacKinnon, R. *J. Mol. Biol.* **2003**, *333*, 965–75.
(70) Pineda, A. O.; Chen, Z.-W.; Bah, A.; Garvey, L. C.; Mathews, F. S.; Di Cera, E. *J. Biol. Chem.* **2006**, *281*, 32922–32928.
(71) Malmberg, N. J.; Varma, S.; Jakobsson, E.; Falke, J. J. *Biochemistry* **2004**, *43*, 16320–16328.
(72) Varma, S.; Jakobsson, E. *Biophys. J.* **2007**, *92*, 966–976.
(73) Pfund, D. M.; Darab, J. G.; Fulton, J. L.; Ma, Y. *J. Phys. Chem.* **1994**, *98*, 13102–13107.
(74) Fulton, J. L.; Pfund, D. M.; Wallen, S. L.; Newville, M.; Stern, E. A.; Ma, Y. *J. Chem. Phys.* **1996**, *105*, 2161–2166.
(75) Seward, T. M.; Henderson, C. M. B.; Charnock, J. M.; Driesner, T. *Geochim. Cosmochim. Acta* **1999**, *63*, 2409–2418.
(76) Zavitsas, A. A. *J. Phys. Chem. B* **2005**, *109*, 20636–20640.
(77) Dong, H.; Liu, W.; Doren, D.; Wood, R. *J. Phys. Chem. B* **2006**, *110*, 18504–18514.
(78) Du, H.; Rasaiah, J. C.; Miller, J. D. *J. Phys. Chem. B* **2007**, *111*, 209–217.
(79) Hofmeister, F. *Arch. Exp. Pathol.* **1888**, *24*, 247–260.
(80) Collins, K. D. *Biophys. J.* **1997**, *72*, 65–76.
(81) Kunz, W.; Nostro, P. L.; Ninham, B. W. *Curr. Opin. Colloid Interface Sci.* **2004**, *9*, 1–18.
(82) Collins, K. D.; Neilson, G. W.; Enderby, J. E. *Biophys. Chem.* **2007**, *128*, 95–104.

and K^+ ions partition from liquid water to higher-order coordinations in the binding sites of proteins without any apparent uphill transitions on the free energy surface.

In order to find clues regarding the underlying physics that allows this structural transition in ion coordination, we now examine more carefully two of the above determinants of ion solvation that will most likely undergo modifications as the ion partitions from liquid water into protein binding sites: (a) the outer-shell environment Γ^C, and (b) the chemistry of the ligands, B.

3.1.2. Role of the Outer-Shell Environment. When ions partition from liquid water (l_W) to binding sites in proteins (ξ), the structural and chemical properties of their outer-shell environments change. This should alter ion solvation energies via modifications in two free energy terms, $\Delta G^{AB}_{n,\Gamma^n}(\xi)$ and $\Delta G^B(\xi)$ (or $\Delta\Phi^B(\xi)$), in eq 12.

When ions bind to proteins, which have much lower dielectric constants than water,[83−87] the ions retain only a partial exposure to the high dielectric water. As a result, the strength of the interaction energies between the clusters AB_n and their outer-shell domains Γ^C decrease, consequently making the term $\Delta G^{AB}_{n,\Gamma^n}(\xi)$ less favorable in comparison to $\Delta G^{AB}_{n,\Gamma^n}(l_W)$. Although this will destabilize clusters of all sizes, it is not expected to alter the relative stabilities between different *n*-fold clusters. In contrast, modifications in the desolvation penalty of the ligand molecule $\Delta\Phi^B(\xi)$ will significantly affect the relative stabilities between different ion coordination structures. A decrease in its magnitude will stabilize all complexes, but more importantly, it will stabilize the higher-order coordinations more than the lower-order coordinations. An increase in its magnitude will have the opposite effect.

Consider a scenario in which this desolvation penalty $\Delta\Phi^B(\xi)$ could somehow be eliminated. For the purpose of reference, let us call the environment that results in negligible values of $\Delta\Phi^B(\xi)$ a quasi-liquid (qL) environment, that is, $\Delta\Phi^B(qL) \sim 0$. If such a hypothetical scenario is encountered in liquid water, the data in Table 3 shows that this will lead to more favorable solvation free energies of all ion−water complexes, that is,

$$\Delta G^A_n(qL_W) = \Delta G^A_{n,\Gamma}(g_W) + n\Delta G^{conc}_W + \Delta G^{AW_n}_{n,\Gamma}(qL) < \Delta G^A_n(l_W) \quad (16)$$

Note that in this model scenario, we continue to consider the local concentration of ligands to be of the same order as that found in liquid water. Setting the desolvation penalty to zero will also lead to an increase in the coordination preferences of both Na^+ and K^+ ions. In this analysis, the 6-fold cluster ends up the most stable cluster for Na^+, while the 8-fold cluster becomes the most stable cluster for K^+.

One obvious model environment in which the desolvation penalty takes up a value of zero is where the static dielectric constant ε of the outer-shell environment Γ^C is equal to unity. In this model situation, both terms $\Delta G^{AB}_{n,\Gamma^n}(\varepsilon = 1)$ and $\Delta\Phi^B(\varepsilon = 1)$ will be zero, which will result in each progressively larger cluster to be more stable, and therefore will result in an increase in the coordination preferences of an ion. But this value of dielectric constant does not represent the environments of biological molecules, which have higher dielectric constants.[83−87] So how large can the dielectric constant be before, for example, the desolvation penalty of a water molecule $\Delta\Phi^W(\varepsilon)$ takes up values large enough to make the coordination number preferences drop back to their original low values in liquid water? Separate calculations of $\Delta\Phi^W(\varepsilon)$ and re-estimations of $n_{preferred}$ for progressively increasing values of ε show that coordination preferences remain high until ε takes up a value greater than three (data not shown). Such low values of the static dielectric constants that yield increased coordination preferences still do not quite represent the environments of biological ion-binding sites.[83−87]

Even though such low values of the macroscopic dielectric constant are unlikely descriptors of Γ^C in biological molecules, an alternative mechanism for reducing electrostatic interactions between binding-site ligands and their environment presents itself upon recognition of the local nature of polar ligand (dipole) solvation in dielectric phases. Dipole−dipole interactions decay as $1/r^3$, and as a result, in contrast to charged species, dipoles should receive the majority of their electrostatic stabilization from direct favorable interactions.[88] This can be demonstrated by computing the free energy penalty $\Delta\Phi^W(\xi)$ associated with extracting a water molecule from a "pseudo-liquid-water" environment in which the high dielectric boundary of liquid water has been radially pushed away to a distance beyond, say, 6 Å from the water molecule (see Figure 2). The free energy penalty $\Delta\Phi^B(\xi)$ associated with extracting any dipolar ligand B can therefore be substantially reduced without actually setting the macroscopic dielectric constant of the entire solvation environment to a very low value. $\Delta\Phi^B(\xi)$ can be substantially reduced by simply eliminating direct favorable interactions of the coordinating ligands with everything else but the binding ion.

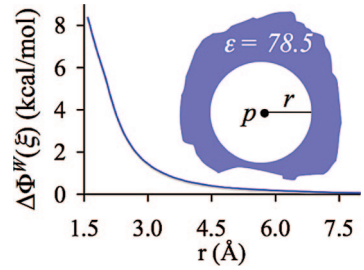


Figure 2. Free energy penalty $\Delta\Phi^W(\xi)$ associated with extracting a water molecule (dipole, *p*) into gas phase from the center of a spherical low dielectric ($\varepsilon = 1$) cavity (radius, *r*) embedded in liquid water ($\varepsilon = 78.5$).

In the case of biological ion-binding sites, the specific conditions that would yield reduced electrostatic penalties for extracting the coordinating ligands for ion complexation consist of:

(a) A proximal environment devoid of only those chemical groups that can directly and favorably interact with the coordinating ligands. For example, in cases where the coordinating ligands are hydroxyl or carbonyl oxygens, the chemical groups that can engage in direct favorable interactions with them are hydrogen bond donors. In proteins, these hydrogen bond donors or competing groups can be supplied either by the protein backbone or by the side chains of some amino acids, including Arg, Lys, His, Cys, Tyr, Thr, Trp, Ser, Gln, and Asn. Their proximity to coordinating ligands in binding sites and also their

(83) Gilson, M. K.; Honig, B. H. *Biopolymers* **1986**, *25*, 2097–2119.
(84) Davis, M. E.; McCammon, A. J. *Chem. Rev.* **1990**, *90*, 509–521.
(85) Simonson, T. *Int. J. Quantum Chem.* **1999**, *73*, 45–57.
(86) Schutz, C. N.; Warshel, A. *Proteins* **2001**, *44*, 400–17.
(87) Hofinger, S.; Simonson, T. *J. Comput. Chem.* **2001**, *22*, 290–305.
(88) Feynman, R. P. *Feynman Lectures On Physics*; Addison Wesley Longman: Reading, PA, 1970; p 1552.

conformational freedom together determine their potency toward favorably interacting with the coordinating ligands and influencing the values of the average desolvation penalties $\langle \Delta\Phi^{B}(\xi) \rangle$. In addition, the values of $\langle \Delta\Phi^{B}(\xi) \rangle$ can also be influenced by the exposure of binding sites to solvent molecules, which can either be water molecules for soluble proteins or a combination of specific lipid headgroups and water molecules for membrane proteins. If the binding sites are exposed to such solvents, the probability to form higher-order coordination structures will be lower, and perhaps that is why all the binding sites listed in Table 1 are mostly buried in the protein matrix. In line with the above argument, the presence of proximal hydrogen bond acceptors can also reduce desolvation penalties by making the ligands more readily accessible to the binding cations.

(b) Specific architectural constraints on the coordinating ligands to prevent them from favorably interacting with any other atoms but the ion. Given the ingrained flexibility in the architectures of many biomolecules, however, it is not straightforward to envision how this strategy can by itself significantly reduce ligand extraction penalties. Nevertheless, we explore its contribution to reducing ligand extraction penalties a little later.

Now consider the opposite scenario in which the desolvation penalty of a ligand molecule, $\Delta\Phi^{B}(\xi)$, is somehow increased, rather than lowered. This will destabilize each and every n-fold ion complex, and in theory could also lead to a reduction in ion coordination preferences. Based on the arguments above, one way to achieve this in experiments is to add certain solutes into water, along with the ions, that will compete with the ions for favorable interactions with the waters. For example, this could occur if the concentration of salt in water is itself increased to values where ions begin to compete with each other for coordination with water molecules. In such a case, an ion will have to "pull" waters away from its neighboring ions to fill its own inner coordination shell, a process that will result in an increase in the desolvation penalty of a water molecule, that is, $\Delta\Phi^{W}(\tilde{l}_W) > \Delta\Phi^{W}(l_W)$, where the tilde over l_W indicates the water solution having the requisite high concentration of ions. This could ultimately lead to a reduction in ion coordination numbers with water, but at the same time it could also lead to a coagulation of ions or an increase in their probability to form cation−anion pairs because, as a pair, they would require wetting by fewer water molecules.

In reality, this indeed happens,[5,7] and from this perspective it can be clubbed together into a more general phenomenon popularly referred to as the Hofmeister effect.[79−82] Analogous to the process in which adding salt coagulates proteins in solution, the process for which this phenomenon is most famous, here salt is "salting-out" itself under conditions of high concentrations. Recent molecular dynamics simulations[78] of Rb^+ and Cs^+ ions in water also show that as their concentrations are increased from millimolar levels to a few molars, their average ion−water coordination numbers drop from 7.5 and 9 to ~6.

Together, the analysis above provides the first clue to explaining how Na^+ and K^+ ions are able to partition readily from aqueous phase into the spatial regions in proteins having much higher coordination numbers, such as those listed in Table 1. We find that increased accessibility of ions to ligands not only enhances the individual stability of each n-fold complex, especially the higher-order complexes, but can also lead to an increase in ion coordination preferences. A schematic of this idea is illustrated in Figure 3. The binding sites in such proteins are, however, not always made up of hydroxyl ligands,[1,2,13−19]

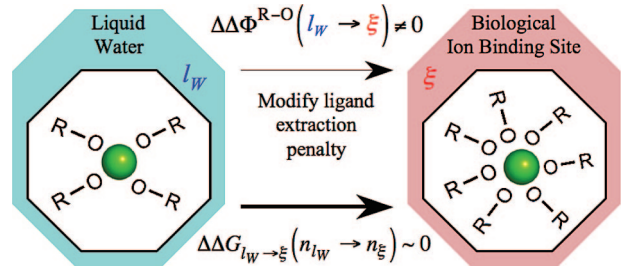


Figure 3. Schematic illustrating the general idea that when an ion is transferred from one environment (say liquid water l_W) to another (say biological ion-binding site ξ), then the free energy cost $\Delta\Delta G_{l_W \rightarrow \xi}(n_{l_W} \rightarrow n_\xi)$ associated with changes in the coordination number of the ion ($n_{l_W} \rightarrow n_\xi$) can be overcome through modifications in free energy penalties $_{\Delta\Delta\Phi^{R-O}}(l_W \rightarrow \xi)$ associated with extracting ligands from the their environments.

such as those provided by water molecules. This raises the question of what role ligand chemistry plays in modulating ion complexation energies.

3.1.3. Role of Ligand Chemistry. From eq 12 and from the data in Table 3, we can see that ligand chemistry will primarily affect ion solvation energies $\Delta G^{A}_{n}(\xi)$ by modulating the values of the two terms, $\Delta G^{B}(\xi)$ and $\Delta G^{A}_{n,\Gamma}(g_B)$. A ligand B that has an electric dipole moment and/or polarizability stronger than water will bind to the ion more tightly, which will make the enthalpic components $\Delta H^{A}_{n,\Gamma}(g_B)$ of cluster free energies more favorable. Concurrently, the same ligand will be repelled electrostatically more by the other ligands in the cluster. This will in turn reduce the stabilities of the clusters ($n \neq 1$). In addition, a ligand with a stronger dipole moment will also have a higher free energy penalty, $\Delta\Phi^{B}(\xi) = -\Delta G^{B}(\xi)$, associated with extracting it from a given environment. This will also decrease the stability of the ion−ligand complex. The overall contribution of ligand chemistry to ion binding will therefore emerge from a balance between at least these three competing physical effects.

The analyses in the previous sections were carried out using water molecules, which are an acceptable representative for the hydroxyl ligands found in biological cation-binding sites. Nevertheless, we supplement the study of the effect of hydroxyl ligands on ion binding by carrying out a similar quasi-chemical analysis of ion solvation in liquid methanol, which has a static dielectric constant of 33. Methanol molecules have an electric dipole moment of 1.7 Debye,[89] which is similar to that of water (1.85 Debye), but more importantly they have a neighboring methyl group that makes them a better representative of the hydroxyl groups found in ion-binding sites in proteins. In numerous cases,[1,2,13−19] including those listed in Table 1, Na^+ and K^+ binding sites also contain carbonyl ligands. In order to understand how carbonyl ligands differ from hydroxyl ligands (water and methanol) with respect to their contributions to ion binding, we carry out a quasi-chemical analysis of the solvation of Na^+ and K^+ ions in liquid formamide. We choose formamide (NH_2CHO) because it has a gas-phase electric dipole moment of 3.73 Debye,[89] which is almost twice as strong as that of a water molecule. Furthermore, formamide under standard conditions exists as a liquid with a static dielectric constant of 109.5, which is similar to that of liquid water.

Table 4 lists the gas phase enthalpies $\Delta H^{A}_{n,\Gamma}(g_B)$ for complexation of Na^+ and K^+ ions with methanol and formamide

(89) Nelson, R. D.; Lide, D. R.; Maryott, A. A. Selected Values of Electric Dipole Moments for Moelcules in the Gas Phase. In *NSRDS-NBS*; 1967, Vol. 10, p 49.

Table 4. Intrinsic Enthalpies (in kcal/mol) of Complexation of Na^+ and K^+ Ions with Methanol (CH_3OH) and Formamide (NH_2CHO) Molecules, Presented Relative to Their Corresponding Values with Water Molecules

	$\Delta\Delta H^{A}_{n,r}(g_W \rightarrow g_B) = \Delta H^{A}_{n,r}(g_B) - \Delta H^{A}_{n,r}(g_W)$			
	$B \equiv CH_3OH$		$B \equiv NH_2CHO$	
n	$A \equiv Na^+$	$A \equiv K^+$	$A \equiv Na^+$	$A \equiv K^+$
1	−2.1	−1.3	−11.5	−9.3
2	−3.4	−2.1	−19.0	−15.4
3	−3.9	−2.6	−21.0	−18.3
4	−3.9	−2.7	−20.0	−18.7
5	−3.9	−3.3	−19.4	−25.3
6	−5.0	−4.4	−14.8	−25.6
7		−5.4		−25.6
8		−5.1		−24.3

molecules. We find that for each cluster size, the ion−methanol and ion−formamide complexes are more stable than their corresponding ion−water complexes. This does not appear surprising for formamide, as it has a stronger dipole moment than water. In the case of methanol, however, this occurs despite the weaker gas-phase dipole moment of methanol, suggesting strongly that methanol is more polarizable than water. Nevertheless, more pertinent to this study is the observation that a ligand that has an electric dipole moment and/or polarizability stronger than water binds to Na^+ and K^+ ions more tightly than water, regardless of the relevant coordination numbers ($n \in \{1, 2, ..., 8\}$). This is contrary to our earlier expectations that increasing the electric dipole moment could destabilize the overall complexation energy via increased ligand−ligand repulsion. The effect of increased ligand−ligand repulsion is certainly visible in larger ion−formamide clusters, but it is not sufficient to override the stability achieved via increased ion−ligand attraction.

The finding above identifies an important role of ligand chemistry in stabilizing ions through short-range interactions within ion−ligand complexes, which has seldom been considered in the context of ion binding to proteins. As can be seen from the data in Table 3, apart from the fact that both Na^+ and K^+ ions receive a significant degree of their stability in liquid water from direct interactions with water molecules in their inner-shell domain, they also receive a significant degree of their stability from interactions with the distant high dielectric environment. This is primarily a consequence of the long-range nature of electrostatic interactions. When ions bind to proteins, which have a much lower dielectric constant than water,[83−87] the ions retain only a partial exposure to the high dielectric water and, therefore, receive reduced stability from the distant environment. Since both carbonyl ligands from formamide and the hydroxyl ligands from methanol bind ions more tightly than water, their presence in biological ion-binding sites can help compensate for the loss associated with the reduced exposure of ions to the high dielectric environment. In line with this observation, higher coordination numbers in the binding sites as compared to those found in water can also serve the same purpose; however, such binding sites need to be accompanied with a local quasi-liquid environment for reduced penalties for ligand extraction.

Ligand chemistry also dictates the thermodynamic cost of extracting ligands from a given environment for ion coordination. Recall that implicit solvent calculations yielded $\Delta\Phi^{W}(l_W) = 8.4$ kcal/mol as the electrostatic penalty to extract a water molecule ligand from bulk water (see data in Table 3). The same calculations show that the electrostatic penalty to extract a methanol molecule from bulk water is 6.2 kcal/mol, while the

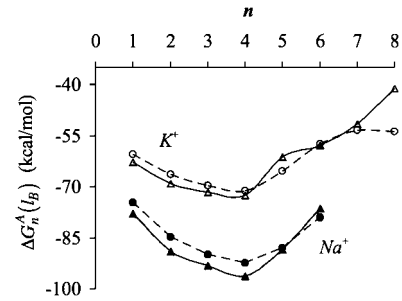


Figure 4. Solvation free energies of $\Delta G^{A}_{n}(l_B)$ of ions $A \in \{Na^+, K^+\}$ in liquid formamide ($l_B \equiv l_{formamide}$, triangles connected using solid lines) and liquid methanol ($l_B \equiv l_{methanol}$, circles connected using broken lines) estimated using the quasi-chemical framework.

electrostatic penalty to extract formamide molecule from bulk water is 10.7 kcal/mol. As expected, the stronger the dipole moment, the higher the desolvation penalty of the ligand. A realization of this aspect of ion solvation can be useful in explaining certain observations, such as in the selectivity filters of NaK channels, where ligands with a weaker dipole moment (hydroxyl) preferentially replace ligands with a stronger dipole moment (carbonyl).[90]

Figure 4 illustrates the solvation free energies of Na^+ and K^+ ions in liquid methanol and in liquid formamide. The radii chosen to define the boundaries of the inner-shell domains are identical to those used for the calculations in liquid water, that is 2.6 Å for Na^+ and 3.1 Å for K^+. We see that ligand chemistry, which had altered the gas-phase enthalpies, does not change ion coordination preferences in these liquids, as both ions still prefer to coordinate with four ligands. In addition, setting the desolvation penalties to zero results in increased coordination preferences, identical to those found in the case of water (data not shown).

Together, this analysis highlights the importance of ligand chemistry in ion binding; however, it does not provide any clues with respect to the ability of ions to partition from their low coordinations in water to overcoordinated binding sites in proteins without energetic barriers.

3.2. Classical MD Analysis. The analysis using the quasi-chemical framework suggested one explanation for how Na^+ and K^+ ions readily partition from aqueous phase into the spatial regions in proteins having much higher coordination numbers. If the accessibility of the ion to ligands is increased via a substantial decrease in the desolvation penalty of ligands ($\langle\Delta\Phi^{B}(\xi)\rangle \rightarrow 0$), then the individual stabilities of the higher-order complexes can be increased, even to values that increase ion coordination preferences. For the sake of reference, we collectively termed the conditions that yielded negligible penalties for ligand extraction a "quasi-liquid" environment. Given the physical nature of interaction of dipoles with their environments (as illustrated in Figure 2), we noted that this penalty could be substantially reduced if direct, favorable interactions of the coordinating ligands with all atoms, except the ion, were eliminated. We identified two different mechanisms to eliminate such direct, favorable interactions: (a) presence of a proximal environment around the coordinating ligands devoid of only those chemical groups that can directly and favorably interact with the coordinating ligands, and (b) specific architectural constraints on the coordinating ligands that prevented them from favorably interacting with any other atoms

(90) Varma, S.; Sabo, D.; Rempe, S. B. *J. Mol. Biol.* **2008**, *376*, 13–22.

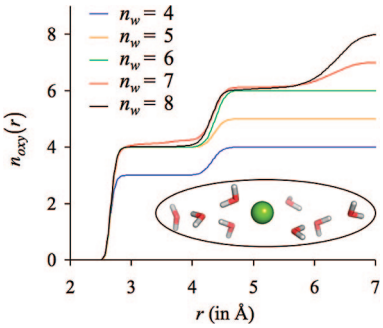


Figure 5. Number distributions $n_{oxy}(r)$ of water oxygens around K^+. Each curve represents an individual simulation in which the natural partial charges of only a subset of randomly selected water molecules n_w in the simulation box are turned on, while the natural partial charges on the remaining water molecules in the simulation box are turned off. The inset illustrates a representative configuration when the partial charges on exactly $n_w = 8$ water molecules are turned on.

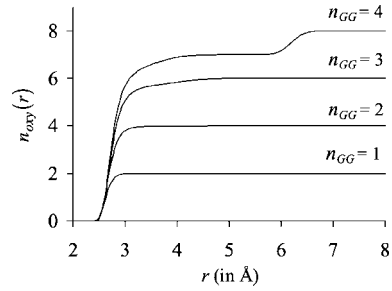


Figure 6. Number distributions $n_{oxy}(r)$ of the oxygens belonging to the glycine dipeptide (*GG*) molecules around K^+. Each curve represents an individual simulation in which n_{GG} glycine dipeptide molecules are present in the simulation box. Note that glycine dipeptide molecules are bidentate ligands, and in contrast to water molecules, provide two oxygen ligands for ion coordination.

but the ion. In order to reaffirm and further explore these ideas, we now switch to classical MD simulations and study the various factors that control the coordination numbers of K^+ ions.

If the condition where the proximal environment is devoid of chemical groups that can directly and favorably interact with the coordinating ligands is by itself sufficient to allow ion partitioning from low coordinations in aqueous phase to overcoordinated states in protein binding sites, then the following test case should hold. If a K^+ ion is introduced in a box of water in a state of overcoordination by, for example, 8 water oxygens, and the partial charges on all the water molecules that are not directly coordinating the ion are turned off (no hydrogen bond donors), then the ion should remain in its original state of overcoordination indefinitely. Note that this does not happen when all the waters in the simulation box carry their natural partial charges.

To test this, we introduce a K^+ ion in a cubical box (length, 30 Å) containing 893 water molecules. We turn off the partial charges on all but a subset of randomly selected water molecules n_w, and generate 4 ns long MD trajectories with periodic boundaries under NVT conditions. Data from the final 1 ns of each trajectory is analyzed and the results of 5 such simulations, where $n_w \in \{4, 5, ..., 8\}$, are illustrated in Figure 5. As expected, the number distributions $n_{oxy}(r)$ around K^+ of the oxygen atoms that carry their natural partial charges change with the value of n_w. Interestingly, however, not more than four such oxygens on average are found to coordinate K^+ directly despite availability of additional water oxygens that carry their natural partial charges. The remaining n_w water molecules prefer to occupy the outer-shell domain Γ^C of K^+, where they line up behind the water molecules that occupy the inner-shell domain. Their dipoles are on average aligned with the electric field of the ion while their oxygens are favorably interacting with the hydrogen atoms of the water molecules occupying the inner-shell domain. This is discernible from the discrete jumps in the $n_{oxy}(r)$ profiles and can also be seen in a representative snapshot of a simulation in Figure 5.

The reason that overcoordination is not achieved in the test scenario above is because lower coordinations are energetically preferred over higher coordinations when the water molecules have the option to drop out of the inner-shell, which is essentially what happens in gas phase, as noted in a previous quantum chemical (MP2) study.[67] This suggests that to achieve overcoordination, a specific variety of rigidity on the coordinating ligands is also required; an architectural constraint that prevents the ligands from dropping out of the inner shell. In the overcoordinated binding sites in proteins, this role invariably needs to be taken up by the protein matrix.

One strategy to specifically test the role of architectural rigidity on ligands is to tether the ligands together, which may prevent them from dropping out of the inner-shell and lining up one behind the other. For this, we choose bidentate ligands in the form of glycine dipeptide (*GG*) molecules. Each of these bidentate ligands provides two oxygens for ligation, instead of the one oxygen supplied by a monodentate water molecule. We retain the simulation protocol used in the previous study, except that in these simulations appropriate numbers of water molecules are substituted with $n_{GG} \in \{1, ..., 4\}$ *GG* molecules, and the partial charges on all the remaining water molecules are turned off. The results of these simulations are illustrated in Figure 6. We now see that K^+ coordinates directly with more than four *GG* oxygens. Nevertheless, the discrete jumps observed in the previous $n_{oxy}(r)$ profiles, which indicate that some ligands have dropped out of the inner-coordination shell, are still present in simulations where $n_{GG} = 3$ or $n_{GG} = 4$. An inspection of the trajectories reveals that the ligands that drop out of the inner-shell are those which hydrogen bond with the amino groups of the adjacent *GG* molecules. In other words, instead of coordinating the K^+ ion, some oxygen atoms of *GG* molecules engage in hydrogen bond interactions with the amino groups of the neighboring *GG* molecules. Essentially, even though there are no water molecules in the simulation box that carry their natural partial charges and can reduce the ion coordination numbers by introducing penalties for ligand extraction, the *GG* molecules can themselves introduce penalties for ligand extraction by supplying hydrogen bond donors in the form of amino groups. Turning off such inter-*GG* hydrogen bond interactions and regenerating the MD trajectories eliminates the discrete jumps in the number distribution profiles, and increases the coordination numbers of K^+ ions (Figure 7a).

Now what happens if, instead of turning off the inter-*GG* hydrogen bond interactions, the density of hydrogen bond donors in Γ^C is further increased? In line with our findings from the quasi-chemical analysis, this should further decrease the number of oxygen atoms directly coordinating the K^+ ion. We test this by reintroducing the natural partial charges on some randomly selected water molecules in the simulation box. As illustrated in Figure 7b, increased population of hydrogen-bond donors does indeed decrease the average numbers of oxygens directly coordinating the K^+ ion. Another point to observe from Figure 7b is that in one of the simulations in which the natural partial charges on 0.009% of water molecules were turned on, one

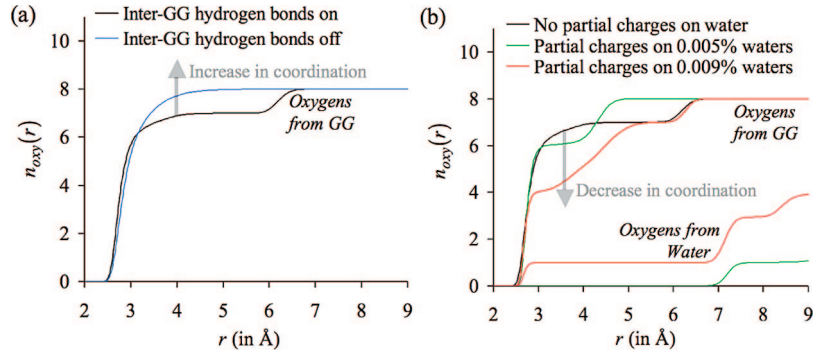


Figure 7. Effect of altering the densities of hydrogen bond donors in Γ^C on the number distributions of oxygens $n_{oxy}(r)$ around K^+. The final configuration of the simulation with four glycine dipeptide (*GG*) molecules in the simulation box ($n_{GG} = 4$ case in Figure 6) was used as the starting configuration for two separate sets of simulations: one in which the densities of hydrogen bond donors were decreased, and another in which the densities of hydrogen bond donors were increased. (a) Since the oxygen atoms of *GG* molecules showed a tendency to hydrogen bond with the amino groups in adjacent *GG* molecules, the densities of hydrogen bond donors were effectively reduced by turning-off all inter-*GG* hydrogen bond interactions. This increased the number of oxygens directly coordinating the K^+ ion, as indicated by a transparent arrow. (b) The densities of hydrogen bond donors were increased by turning on the natural partial charges of a certain percentage of water molecules in the simulation box. This decreased the total number of oxygen atoms directly coordinating the K^+ ion, as also indicated by a transparent arrow. Note that in the case where the partial charges on 0.009% of the waters are turned on, this reduction in coordination number occurs despite one oxygen atom from water substituting for the oxygen atoms from the *GG* molecule.

water molecule directly coordinates the K^+ ion, preferentially replacing a carbonyl oxygen atom from the *GG* molecules. This is an example of a situation described in the previous section where a ligand with a weaker dipole moment (hydroxyl) preferentially replaces a ligand with a stronger dipole moment (carbonyl), primarily because the desolvation penalty associated with extracting a ligand with a weaker dipole moment is smaller.

In the analysis above, ion coordination properties were studied under conditions of ambient temperature (298.15 K). In contrast, the X-ray data of biomolecules in Table 1, where ions were found to be overcoordinated with respect to ion coordinations in liquid water, was collected after flash freezing the crystals to very low temperatures below 200 K. Is it legitimate to compare our findings at physiological temperatures to those obtained under conditions of very low temperatures? In other words, does flash freezing of crystals introduce any artifacts in the ion coordination structures? It is common knowledge that temperature manifests itself as fluctuations in interacting particles, so from that standpoint it seems that the process of flash freezing should basically dampen thermal fluctuations and "trap" the ion coordination structures in their minimum energy configurations. Despite the fact that temperature does not alter potential energy surfaces (PES), it can alter transition probabilities across energy barriers on the PES, ultimately modifying the overall topologies of the ion coordination structures. No hard-and-fast rules, however, can predict the temperature-response of electrostatically interacting particles, as each such system will have its unique PES.[91] Nevertheless, we test this for our candidate 8-fold K^+−*GG* complex.

We start with the final snapshot of the trajectory of a K^+ ion generated with four *GG* molecules in the absence of hydrogen bond donors (including inter-*GG* hydrogen bond interactions), and generate new 4 ns long trajectories at lower temperatures of 200 and 100 K. The results of these simulations are illustrated in Figure 8, where they are also compared to the results of the simulation at ambient temperature (298.15 K). Despite the observation that the volume of coordination shell contracts, none of the ligands drop out into the second shell of the ion. In addition, the average ion coordination structures at lower temperatures remain similar to the average ion coordination structure at ambient temperature. In other words, we find that lowering the temperature did not alter the overall topology of this 8-fold complex, and congruous with our expectations, lowering the temperature reduced thermal fluctuations and in effect "tightened" this coordination complex. Furthermore, when the temperature was lowered, the average K^+−O distances decreased only marginally. At a temperature of 298.15 K, the average K^+−O distance was 2.90 ± 0.06 Å. When the temperature was lowered to 200 K, the average K^+−O distance

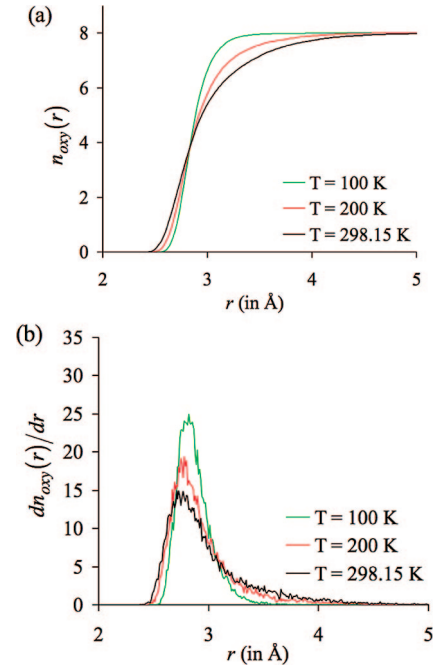


Figure 8. Effect of temperature on (a) the number distributions of oxygen atoms $n_{oxy}(r)$ and (b) their derivatives $dn_{oxy}(\mathbf{r})/dr$ around K^+. The final snapshot of the simulation of K^+ at a temperature $T = 298.15$ K with four *GG* molecules embedded in a simulation box without any hydrogen bond donors (including inter-*GG* hydrogen bond interactions) was used as the starting configuration for simulations at lower temperatures. Lowering the temperature "tightened" this coordination complex, without altering its overall topology. Note that, while the volume of coordination shell contracts, none of the ligands drop out into the second shell of the ion.

(91) Wales, D., *Energy Landscapes: With Applications to Clusters, Biomolecules and Glasses*; Cambridge University Press: New York, 2003.

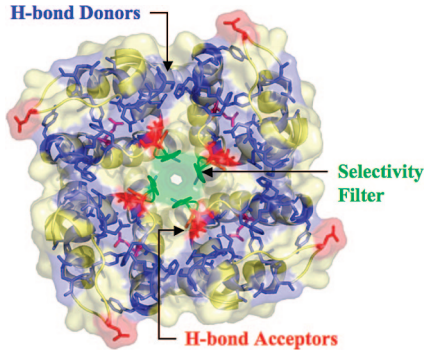


Figure 9. Distribution of hydrogen bond donor and acceptor side chains in a representative structure of potassium channels, KirBac1.1 (PDB ID: 1P7B). The orientation of the channel is as viewed from its extra-cellular end. The selectivity filter, which provides the backbone carbonyl oxygens to form 8-fold K^+ binding sites, is colored green. The side chains of amino acids H, R, K, N, Q, T, Y, W, S, and C that contain hydrogen bond donors and do not belong to the selectivity filter are illustrated as sticks and colored blue. The side chains of amino acids D and E that contain hydrogen bond acceptors are colored red. In the crystallographic configuration, all relevant side-chain hydrogen bond donor groups are at a distance of at least 6.8 Å from the coordinating carbonyl oxygens of the selectivity filter.

decreased to 2.86 ± 0.05 Å, and when the temperature was further lowered to 100 K, the average K^+-O distance decreased to 2.82 ± 0.03 Å. This strongly suggests that the average ion−ligand distances collected at low temperatures are likely to be only slightly smaller than their values at physiological temperature, and it is indeed legitimate to compare our findings at physiological temperatures to those obtained under conditions of very low temperature.

Together, the classical MD analysis reaffirms the finding from quasi-chemical analysis that ion coordination preferences can be altered via changes in the penalty associated with extracting ligands from their solvation environments. This penalty can be altered not only via specific modifications in the local environments of the coordinating ligands, but also via modifications in the conformational flexibility (or rigidity) of the coordinating ligands.

3.3. Environmental Analysis of Overcoordinated Binding Sites in Biomolecules. It is straightforward to see that the local structural and chemical factors that lead to a substantial reduction in the desolvation penalties of ligands (or lead to the formation of local quasi-liquid pockets around the ligands) and enable formation of overcoordinated structures without energetic barriers will be unique to each biological ion-binding setting. In this section, we undertake a structural and chemical analysis of the environments of all the overcoordinated binding sites listed in Table 1.

We first identify indications of the presence of a local quasi-liquid environment around the binding sites of potassium channel selectivity filters, where K^+ ions are seen to coordinate directly with 8 ligands in the X-ray structure.[20−23] First, the concentration of ligands in the apo binding site is ∼23 M, which is comparable to the concentration of water ligands in aqueous phase. This was estimated by dividing the total number of carbonyl ligands available in the selectivity filter of the apo state of potassium channels[69] by the volume of the selectivity filter. Second, there are no relevant side-chain hydrogen bond donors within the local environments of the ligands in any of the four known structures[20−23] of potassium channels (KcsA, MthK, KvAP, and KirBac1.1), as can be seen in an illustration of KirBac1.1 in Figure 9. Third, all the proximal backbone hydrogen bond donors are hydrogen bonded to other functional groups in the protein and involved in maintaining the integrity of the protein's three-dimensional fold. Fourth, there are some hydrogen bond acceptor groups present in the vicinity of the filter in the form of Asp and Glu residues. The negative charges on these residues will serve to prevent the carbonyl ligands from turning away from the ion-binding sites. And finally, the packing of residues and the hydrogen bond network behind the selectivity filter also provide a certain degree of structural rigidity to the ligands themselves, as supported by the finding that deformation of the filter structure requires an appreciable energetic cost.[92] These observations indicate how potassium channels can sustain K^+ ions bound in states of high coordination.

Table 5. Normalized Number Densities of Hydrogen Bond Donors around the Ligands in the Ion-Binding Sites of Proteins listed in Table 1[a]

biomolecule	$\eta_{\text{H-donor}}/\eta_w \times 100$ ($r_d \in (6,12]$)	$\eta_{\text{H-donor}}/\eta_w \times 100$ ($r_d \in [0,6]$)	$\eta_{\text{free-H-donor}}/\eta_w \times 100$ ($r_d \in [0,6]$)
pyruvate kinase (1A49)	6.90	6.17	0.00
chaperone (1HPM)	6.58	1.30	0.47
lysozyme (1V7T)	8.07	1.63	0.00
neurotransmitter transporter (2A65)	4.84	2.22	0.26
dialkylglycine decarboxylase (2DKB)	2.65	0.52	0.26
thrombokinase (2BOK)	2.95	0.98	0.33

[a] The normalization factor η_w is the number density of bulk water. For a given binding site, individual number densities were first estimated for all of its ligands, and then averaged. These average number densities were computed for two different spherical subvolumes around the ligand, one limited to a region within 6 Å from the ligand, while the other extends from 6 to 12 Å from the ligand. In the estimation of $\eta_{\text{H-donor}}$, hydrogen bond donors from all side-chain residues were considered, including Arg, Lys, His, Cys, Tyr, Thr, Trp, Ser, Gln, and Asn, while in the estimation of $\eta_{\text{free-H-donor}}$, only those hydrogen bond donors were considered that were not engaged in other favorable interactions, including salt-bridges and hydrogen bonds.

Table 5 presents a survey of the environments of the coordinating ligands in the remaining ion-binding sites of Table 1. Columns 2 and 3 list the number densities $\eta_{\text{H-donor}}$ of side-chain hydrogen bond donors in the distant ($r_d \in (6,12]$) and local ($r_d \in [0,6]$) environments of the coordinating ligands, where r_d is the distance in Å from the ligand and not from the ion. Barring the case of pyruvate kinase, the local environments carry far fewer hydrogen bond donors than found in the distant environment. Irrespective, the densities of hydrogen bond donors in the local environments are almost 2 orders of magnitude smaller than the density of hydrogen bond donors offered by liquid water. On the basis of the analysis above, this will reduce the penalties associated with extracting these ligands for ion coordination, and increase the stability of ions in these proteins. Now, if only those hydrogen bond donors are considered that are not engaged in other favorable interactions, such as salt-bridges and hydrogen bonds with other groups in the proteins, then we find that the number densities $\eta_{\text{free-H-donor}}$ drop to negligible values. In other words, hardly any freely available side-chain hydrogen bond donors are present in the local neighborhoods of the coordinating ligands that can engage in favorable interactions with them and disrupt the formation of higher-order ion−ligand complexes.

In the local neighborhood of the ligands, the nonzero values of $\eta_{H-donor}$ are, in fact, expected to play a special role during ion binding. In the absence of bound ions, the coordinating ligands in the binding site will experience large electrostatic repulsions from each other. On the one hand, in a scenario where

(92) Miloshevsky, G. V.; Jordan, P. C. *Biophys. J.* **2008**, In press.

proximal hydrogen bond donors are absent, the binding site will build up significant strain energy and may have to undergo large structural changes. On the other hand, if proximal hydrogen bond donors are present, then this strain can be relieved through favorable interactions, as has been seen recently in the case of cytosolic phospholipases.[72] This brings us to believe that, although the presence of proximal hydrogen bond donors can disrupt formation of higher-order complexes, they can at the same time relieve the strain energy that builds up in the absence of bound cations without need for large structural changes. Perhaps there exists an optimum density of proximal hydrogen bond donors that benefits both aspects of the ion-binding process, with its implication being tightly intertwined with the individual functions of ion-binding sites. We note that this analysis is based on static X-ray structures, and an analysis using Monte Carlo sampling or MD simulations is expected to provide a better perspective on the distributions and the dual roles of proximal hydrogen bond donors in ion binding, especially those belonging to the protein backbone that were not considered during analysis other than in the case of potassium channels.

4. Conclusions and Summary

In this work, we focused our efforts to understand one specific aspect of the ion-binding phenomenon in biomolecules. When ions, such as Na^+ and K^+, dissolve in liquid water, they exhibit strong preferences toward the numbers of ligands inside their inner-coordination shells. Alternatively phrased, there is a free energetic cost associated with transferring ions from their preferred coordination states in water to states having different coordination numbers. Such free energetic costs become exceedingly large if transitions are considered to coordination numbers that are very different from preferred values. Nevertheless, there are several biomolecules[1,2,13−19] that, based primarily on X-ray data, appear to bind these ions using coordinations that are much higher than those preferred in liquid water. Why do not energetic costs surface during such structural transitions from water to overcoordinated binding sites (relative to water) in biomolecules? What deflates these energetic costs and allows ions to partition readily back-and-forth between liquid water and the binding sites in biomolecules? What physical principles explain such structural transitions in ion coordinations?

We utilized a three-way strategy to first expand our understanding of the essential nature of ion solvation in liquid media and then determine the underlying physics that explains this ion coordination anomaly. First, quantum chemical simulations within the statistical framework of the quasi-chemical theory of solutions[24−27,29,30] was invoked to understand the individual and combined roles of the various factors that drive ion solvation in liquid media. Note that the goal here was not to determine the exact contributions arising from the various factors that drive ion solvation but only to assess their participation in a qualitative manner. Then classical MD simulations were carried out to probe the inferences derived from the quasi-chemical analysis. Finally, the findings from the quasi-chemical and MD analyses were examined in the context of the architectures and chemical compositions of the overcoordinated binding sites in biomolecules.

This analysis revealed an important ingredient to the recipe for lowering the energetic costs associated with partitioning ions from liquid water into any given overcoordinated binding site. If the penalty associated with extracting ligands from their solvation environments to form ion coordination complexes is substantially reduced, that is $\Delta\Phi^B(\xi) \sim 0$, then the stabilities of higher-order complexes are dramatically increased, even to extents that can drive up the coordination preferences of ions. For the sake of reference, we collectively term the conditions that yield negligible penalties for ligand extraction a "quasi-liquid" environment. Given the physical nature of interaction of ligands (dipoles) with their environments (as illustrated by Figure 2), this penalty can be substantially reduced by simply eliminating direct favorable interactions of the coordinating ligands with all atoms except the ion. Biomolecules can utilize two different strategies to accomplish this:

(a) Tune the local environments of their coordinating ligands: The local environments of the coordinating ligands can be tuned such that the local number densities of competing groups that can favorably interact with the coordinating ligands are either significantly reduced or made exceedingly less accessible for interaction through specific steric and electrostatic interactions in the protein matrix.

(b) Impose specific architectural constraints on the ligands themselves: Specific architectural constraints can be imposed on the coordinating ligands to prevent them from favorably interacting with any other atoms but the ion. Given the ingrained flexibility in the architectures of biomolecules, however, it is not straightforward to envision how this strategy can by itself significantly reduce ligand extraction penalties.

If a substantial reduction in the local densities of competing groups, which can favorably interact with the coordinating ligands, can increase ion coordination preferences, then a substantial increase in the number densities of such competing groups should lead to a drop in coordination preferences. This is indeed true, as was demonstrated using classical MD simulations. In fact, this is reminiscent of the effect of salt concentrations on an ion's coordination number. When the concentration of salt in water is increased to values where ions begin to compete with each other for coordination with water molecules, the ion−water coordination numbers decrease. This has been observed both experimentally[5,7] and theoretically.[78] It is also tempting to compare this phenomenon to the famous Hofmeister effect.[79−81] Analogous to the process in which adding salt precipitates proteins, the process for which the Hofmeister effect is most famous, here salt is "salting-out" itself under conditions of high concentrations.

The complete absence of competing groups in the local neighborhoods of binding sites will maximize the stabilities of the higher-order complexes. In such a scenario, however, when there are no ions present in the binding sites, the coordinating ligands in the binding site will experience large electrostatic repulsions from each other. As a result, the binding site will build up significant strain energy and may have to undergo large structural changes. In contrast, if proximal competing groups are present, then this strain can be relieved through favorable interactions, as has been seen recently in the case of cytosolic phospholipases.[72] This brings us to believe that although the presence of proximal competing groups can disrupt formation of higher-order complexes, they can at the same time relieve the strain energy that builds up in the absence of bound ions without need for large structural changes. The advantage of preventing the binding site from undergoing large structural changes in the absence of ions is that it helps maintain high ligand concentrations, which enhances ion-binding energies. In theory, there should exist an optimum density of proximal competing groups that benefits both aspects of the overall ion-binding process. What are these optimum densities? Has evolution indeed carved out such optimum densities in biomolecules to boost their function? Can we harness them to engineer

desirable environmental responses? Answers to such questions require a more extensive study of this phenomenon, which we anticipate will be the focus of future studies.

Our studies also show that modifying ligand chemistries needs not alter ion coordination preferences and/or ion solvation energies. On the one hand, ligands that have larger dipole moments or polarizabilities (field strengths) are energetically more expensive to extract from their solvation environment for ion coordination. This will serve to decrease the stabilities of coordination complexes. On the other hand, ligands that have stronger field strengths will bind more strongly to the ion, which will increase the stabilities of coordination complexes. It is easy to see that these competing effects can on certain occasions cancel each other out and yield unchanged ion coordination preferences and/or ion solvation energies. This we argue is the primary reason why ion solvation structures or energies need not differ between solvents that have vastly different dielectric properties. For example, the solvation energies of Na^+ and K^+ ions in liquid water are similar to their corresponding values in liquid formamide, despite the fact that formamide has a dipole moment twice as large as that of water.

Investigations concerning the effect of temperature on ion coordination led to the finding that lowering the temperature "tightens" coordination complexes, at least for the particular case considered in this study. When MD simulations were carried out at lower temperatures of 100 and 200 K, temperatures to which crystals are flash frozen for use in structure determination by X-ray studies, the thermal fluctuations decreased without altering the overall topology of the coordination complex. In addition, the average ion−ligand coordination distances were minimally affected. This strongly suggests that (a) the average ion−ligand distances collected at low temperatures are likely to be good representatives of their values at physiological temperature, and (b) it is indeed legitimate to compare our findings at physiological temperatures to those obtained under conditions of very low temperature.

Our overall findings testify to the intricate designs of the overcoordinated binding sites in biomolecules. So why had evolution taken the trouble to design such binding sites? Are they mere remnants of the act of serendipity, or do they serve any purpose? We identify two important roles of overcoordination in biomolecules:

(a) Overcoordinating ions can increase ion stability in special cases where ions are required to bind within the interiors of biomolecules, including the narrow passageways of ion channels. When ions bind to buried-sites in biomolecules, overcoordination can serve to compensate for the loss in exposure of the ion to the high dielectric water. For example, the binding energy of K^+ ion with six or more carbonyl ligands (supplied by three of more glycine dipeptide ligands) is similar to the hydration energy of the K^+ ion.[56] This will allow K^+ ions to readily partition from water into such overcoordinated binding sites without the need for stability from the high dielectric water. It is interesting to note that such an outcome simultaneously requires that the ion-binding site has reduced exposure to water for reduced penalties for ligand extraction.

(b) Overcoordination is a potent mechanism for K^+/Na^+ selectivity[56,90,92−96] in the fluctuating (but not liquid)[56,90,93,94,97] ion-binding sites of potassium channels. Reducing the extent of overcoordination decreases ion selectivity,[56,90,93,94,98] which appears to have been utilized by nature to create different varieties of potassium channels that exhibit weak, intermediate, and strong K^+/Na^+ selectivity.[56,90,99] Certainly, binding ions into states of overcoordination may provide other advantages to biomolecules, and it will be interesting to see what future investigations reveal.

Acknowledgment. We thank Dr. Eric Jakobsson and the NCSA at the University of Illinois Urbana−Champaign for providing us compute time on its SGI Altix shared-memory cluster. This work was supported, in part, by Sandia's LDRD program and, in part, by the National Institutes of Health through the NIH Road Map for Medical Research. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL8500.

Supporting Information Available: Optimized structures of all ion−water complexes and reference 36 with complete list of authors. This material is available free of charge via the Internet at http://pubs.acs.org.

JA803575Y

(93) Bostick, D.; Brooks, C. L., III *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9260–9265.
(94) Thomas, M.; Jayatilaka, D.; Corry, B. *Biophys. J.* **2007**, *93*, 2635–2643.
(95) Vora, T.; Bisset, D.; Chung, S. H. *Biophys. J.* **2008**, *95*, 1600–1611.
(96) Fowler, P. W.; Tai, K.; Sansom, M. S. *Biophys. J.* **2008**, In press.
(97) Asthagiri, D.; Pratt, L. R.; Paulaitis, M. E. *J. Chem. Phys.* **2006**, *125*, 24701–24706.
(98) Noskov, S. Y.; Berneche, S.; Roux, B. *Nature* **2004**, *431*, 830–834.
(99) Shealy, R. T.; Murphy, A. D.; Ramarathnam, R.; Jakobsson, E.; Subramaniam, S. *Biophys. J.* **2003**, *84*, 2929–2942.